In the Claims

Please amend claims 1 and 7 as marked-up in **Exhibit A** attached hereto, by deleting the bracketed matter and inserting the underlined matter. A clean copy of claims 1 and 7, as amended, is attached as **Exhibit B**.

REMARKS

Claims 1-12 are pending in the present application and presented for reconsideration. Applicants have hereinabove amended claims 1 and 7 to place the claims in better form for examination, without narrowing the scope of the claimed invention.

Applicants maintain that no new matter is presented by this amendment. Accordingly, Applicants respectfully request that this Amendment be entered.

Rejection Under 35 U.S.C. §102(a)

On page 2 of the January 27, 2003 Office Action, claims 1-12 were rejected under 35 U.S.C. §102(a) as being anticipated by the National Institute of General Medical Sciences (NIGMS) Protein Structure Initiative (PSI) Meeting Summary dated April 24, 1998 (hereinafter "the NIGMS PSI paper").

The Examiner stated that the NIGMS PSI paper "summarized the discussion of a one-day meeting held on April 24, 1998 to experimentally determine 3D structures of protein families via a representative protein molecule (target) from each family." The Examiner also stated that "the text of the [NIGMS PSI paper] was last updated June 2, 1998."

The Examiner stated that "protein sequences were compared using sequence homology to define families and targets selected." The

Examiner also stated that "the targets are then cloned into plasmids for overexpression, and purified for use in crystallization trials." The Examiner further stated "those targets successfully crystallized have X-ray crystallography and protein structure determination performed."

The Examiner stated that "synchrotrons, multiwavelength anomalous diffraction (MAD), and selenomethionyl enrichment are specifically disclosed." The Examiner also stated that "structural and functional properties would be predicted." The Examiner further stated that "in particular, identification of protein fold motifs is disclosed."

The Examiner stated that "the results of the analysis is put in a database with any additional information that may be helpful for further experiments." The Examiner also stated that "the database is updated and annotated as research progresses." The Examiner further stated that "the database is intended to be accessible to all researchers."

The Examiner stated that "implicit in this document is a system containing the component parts (database and means) for executing each of the steps of the method." The Examiner also stated that "it is noted that while at least inventor Hendrickson is indicated to have been present at the meeting, the document cannot be considered his own work as it represents the recommendations and conclusions of all participants."

Applicants maintain that the claimed invention cannot be anticipated by the NIGMS PSI paper because it has not been shown that the NIGMS PSI paper is a prior art printed publication. A printed publication must be disseminated and publicly available before the invention by applicant in order for the printed

publication to be a §102(a) bar.

A Declaration Under 37 C.F.R. § 1.132 Of Paul Teng is attached hereto as **Exhibit C**. Exhibit 1 attached to the Declaration is a computer printout of an electronic mail message received by Applicants' undersigned attorneys (hereinafter "the e-mail message"). The e-mail message states that it is from Daniel Hogan, and two electronic files are attached thereto. Mr. Hogan is the current Web Site Manager of the NIGMS and is responsible for maintaining the web pages of the NIGMS. Exhibit 2 attached to the Declaration is a computer printout of one of the two electronic files, which is named "protein_structure.html", attached to the e-mail message. Exhibit 3 attached to the Declaration is a computer printout of the other electronic file, which is named "protein_structure.html.bak", attached to the e-mail message.

The e-mail message states that the first electronic file (Exhibit 2 attached to the Declaration) corresponds to the current NIGMS web page (hereinafter "the current NIGMS PSI web page file"), at the web address http://www.nigms.nih.gov/news/reports/protein_structure.html, on which the NIGMS PSI paper is displayed. The e-mail message also states that the other electronic file (Exhibit 3 attached to the Declaration) corresponds to the previous version of the NIGMS web page (hereinafter "the previous NIGMS PSI web page file") on which the NIGMS PSI paper was displayed. The e-mail message further states that (a) the current NIGMS PSI web page file was last modified March 8, 2000, and (b) the previous NIGMS PSI web page file was last modified May 12, 1999. Further, the e-mail states that no other versions of the NIGMS web page on which the NIGMS PSI paper is displayed were found at NIGMS.

Accordingly, the earliest date that can be attributed to the NIGMS PSI paper is May 12, 1999. Since it has not been shown that the NIGMS PSI paper was disseminated and publicly available prior to the January 22, 1999 filing date of the subject application, Applicants maintain that the NIGMS PSI paper is not prior art to the invention claimed in the present application.

In addition, Applicants maintain that the claimed invention cannot be anticipated by the NIGMS PSI paper because the NIGMS PSI paper fails to disclose each and every element of the claimed invention in an enabling manner.

The present claims are directed to a novel and unobvious system (independent claim 1) and method (independent claim 7) for determining experimentally a plurality of three-dimensional protein structures. Sequence information for a first plurality of proteins and structural information and functional information for a second plurality of proteins (which may be a subset of the first plurality) are systematically organized into a database. The sequence information, structural information and functional information stored in the database are used with a bioinformatics tool to cluster the plurality of proteins into families. In each such family, the members have homologous sequences. For each family, selected members of the family are chosen as target proteins (for structure determination). The target proteins are synthesized. Synthesized products are screened, processed and crystallized into specimen crystals. The specimen crystals are tested to determine the specimen crystals that are suitable for diffraction measurements. High-throughput X-ray crystallography is performed on the suitable specimen crystals. Diffraction data obtained from the X-ray crystallography is analyzed and used to build and refine an atomic model of the corresponding target protein. The refined model of the target protein is analyzed

using sequence information corresponding to other family members and structural information corresponding to other proteins, which are stored in the database, to determine functional motifs and surface characteristics.

The claimed system and method may be used to develop a comprehensive structural genomics database. As new structural information and functional information are determined/obtained, the database is appropriately updated using the newly determined/obtained information and bioinformatics tools. For example, the database may be updated using the bioinformatics tool and the developed homology model to link the refined model of the target protein to other databases which store information concerning biological pathways and functional annotation. Such a system would be particularly suitable for a number of applications, such as, for example, drug design, which often requires ready access to the wealth of genomics information.

The NIGMS PSI paper, as understood by Applicants, provides a general discussion of the state of the art in connection with a proposed Protein Structure Initiative which includes a proposed method for determining the structure of at least one member from selected protein families. The proposed method consists of (1) identification of protein families and target selection, (2) generation of protein for biophysical analysis, (3) structure determination of selected targets, and (4) structure and sequence analysis.

The NIGMS PSI paper does not, however, disclose each and every element of the claimed invention set forth in amended claims 1 and 7.

The NIGMS PSI paper does not disclose, for example, (a)

systematically organizing sequence information for a first plurality of proteins, and structural information and functional information for a second plurality of proteins into a database, and (b) using at least one bioinformatics tool and the sequence information, structural information and functional information stored in the database, to cluster the proteins into a plurality of families, in which, for each family, members of the family have corresponding homologous sequences, as provided by the claimed invention recited in amended claims 1 and 7.

The NIGMS PSI paper does not teach the benefit of a comprehensive genomics database of structural information, sequence information and functional information, as may be achieved through the claimed invention. Although the NIGMS PSI paper generally mentions sequence analysis and structure prediction, it does not refer to bioinformatics nor to specific bioinformatics tools which can be used to maximize the utility of a comprehensive genomics database. The NIGMS PSI paper also fails to recognize the benefits of having a comprehensive genomics database from which information that may be readily retrieved for the protein synthesis process, once target proteins are selected. The NIGMS PSI paper simply does not provide an enabling disclosure for a system for determining experimentally a plurality of three-dimensional atomic structures, each of which is associated with a corresponding protein, comprising, amongst other elements, a database of protein sequence information, structural information and functional information and at least one bioinformatics tool adapted to use the protein sequence information, structural information and functional information stored in the database, as provided by the claimed invention recited in amended claim 1.

As another example, although the NIGMS PSI paper does provide a brief discussion of protein crystallization, it fails to disclose

testing specimen crystals of a target protein for predetermined diffraction characteristics to determine the specimen crystals that are suitable for diffraction measurements, as provided by the claimed invention recited in amended claims 1 and 7.

Since the NIGMS PSI paper does not disclose each and every element of the claimed invention, it cannot anticipate the claimed invention.

Accordingly, Applicants respectfully request that the Examiner reconsider and withdraw the rejection of claims 1-12 under 35 U.S.C. §102(a).

Rejection Under 35 U.S.C. §102(a)

On page 3 of the January 27, 2003 Office Action, claims 1-12 were rejected under 35 U.S.C. §102(a) as being anticipated by Gaasterland (Nature Biotechnology, July 1998) (hereinafter "the Gaasterland paper").

The Examiner stated that "Gaasterland reviews the goals and initial results of the structural genomics initiative." The Examiner also stated that "results from the pilot project presented in January 1998 at the Argonne National Laboratory are discussed." The Examiner further stated that "flow diagrams (Figures 1 and 2) and particular bioinformatics tools (Table 1) with the accompanying discussion are considered to disclose the claimed methods and systems."

Applicants maintain that the claimed invention cannot be anticipated by the Gaasterland paper because the Gaasterland paper fails to disclose each and every element of the claimed invention in an enabling manner.

The Gaasterland paper, as understood by Applicants, provides a summary and survey of the state of the art in structure determination and computational methods. Figure 1 of the Gaasterland paper shows a general diagram of tasks for structural genomics. Figure 2 of the Gaasterland paper shows a flow of structural data, as proposed by the paper, for a biological understanding of structures through computational techniques.

Gaasterland paper does not, however, disclose each and every element of the claimed invention set forth in amended claims 1 and 7.

Although the Gaasterland paper mentions a crystallization step in the structure determination process shown in Figure 1 therein, the Gaasterland paper does not disclose, for example, testing the plurality of specimen crystals for predetermined diffraction characteristics to determine the specimen crystals that are suitable for diffraction measurements, as provided by the claimed invention recited in claims 1 and 7.

The Gaasterland paper also fails to disclose other features of the claimed invention, such as the following: (a) freezing the specimen crystals of the target protein, wherein the suitable specimen crystals are frozen before being measured for the diffraction data (claims 2 and 8); (b) cloning for each family cDNAs corresponding to family members into a plurality of expression vectors for a plurality of expressions systems (claims 3 and 9); (c) the high-throughput crystallography is performed using a synchrotron storage ring having undulator beamlines (claims 4 and 10); and (d) selenomethionine is incorporated in the synthesized target proteins synthesized, and the multiwavelength anomalous diffraction phasing method is used to analyze diffraction data measured for selenomethionyl proteins

(claims 5 and 11).

Since the Gaasterland paper does not disclose each and every element of the claimed invention, it cannot anticipate the claimed invention.

Accordingly, Applicants respectfully request that the Examiner reconsider and withdraw the rejection of claims 1-12 under 35 U.S.C. §102(a).

Rejection under 35 U.S.C. §112, first paragraph

On page 4 of the January 27, 2003 Office Action, claims 1-12 were rejected under 35 U.S.C. §112, first paragraph.

The Examiner stated that "the specification, while being enabling for some aspects of the claimed method and system, does not reasonably provide enablement for the breadth of what is encompassed." The Examiner also stated that "the specification does not enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the invention commensurate in scope with these claims."

The Examiner stated that "the claims encompass use of a synchrotron." The Examiner also stated that "such a piece of equipment is not readily available for use or purchase." The Examiner further stated that "beamline time must be applied for at existing synchrotron facilities such as Argonne National Laboratory and that a synchrotron is not readily available for purchase."

The Examiner stated that "Holmes (Philosophical Transactions of the Royal Society of London Biological Sciences, December 1999) which establishes that synchrotrons cost approximately

$100,000,000.00, take approximately 10 years to build, and can't be purchased from a catalog." The Examiner also stated that "Holmes also documents that there are only a limited number of synchrotron facilities in the world."

The Examiner stated that "with respect to clustering sequences into families, Heger et al. (Progress in Biophysics & Molecular Biology, 2000) establishes that even well after the filing date, this was a non-trivial computational problem." The Examiner also stated that "family assignment in structural genomics is specifically discussed at pages 334-335 and that in February 1999 (after the instant filing date) target lists of unknown protein families were still being developed."

The Examiner stated that "the claims encompass any and all proteins and families, yet it would have been well known in the art at the time of the invention that membrane bound, proteins found in complexes, and insoluble proteins were problematic for crystallization and X-ray crystallographic structure determination." The Examiner also stated that "the pilot projects for structural genomics as discussed by Gaasterland were selected in part to determine what classes of proteins were feasible to attempt structure determinations upon and were selected because they were thought to be easier."

The Examiner stated that "to the degree that the system of claims 1-6 is directed to an integrated, turn-key system or fully automated system with the recited database and various recited means, such systems are not enabled." The Examiner also stated that "the examiner is not aware of any integrated system or fully automated system possessing all (or even most) of the recited components at the time of the invention." The Examiner further stated that "even particular aspects of the system were not fully

automated."

The Examiner stated that "it does not appear that crystallization apparatus (to produce suitable crystals from proteins) was integrated to or automated in association with a synchrotron at the time of the invention nor would it have been feasible to do so." The Examiner also stated that "the crystals would have been grown and stored until such time as the synchrotron was available for diffraction analysis." The Examiner further stated that "it is deemed undue experimentation to practice this embodiment of the invention."

In response, without conceding the correctness of the Examiner's position but solely to advance the prosecution of the present application, Applicants have hereinabove amended claims 1 and 7. Applicants maintain that the claim amendments do not narrow the scope of the claimed invention, but rather place the claims in better form for examination.

Applicants respectfully traverse the rejection and maintain that the claimed invention as set forth in the amended claims is adequately supported by the specification.

The patent laws and rules require a disclosure of the claimed invention that enables a person of ordinary skill in the art to practice the claimed invention [see Christianson v. Colt Industries Operating Corp., 822 F.2d 1544, 1562 (Fed.Cir. 1987)]. Improvements made to the described embodiments of the claimed invention after filing of the application do not mean that an enabling disclosure was not made [see Hormone Research Foundation, Inc. v. Genentech, Inc., 904 F.2d 1558, 1568 (Fed.Cir. 1990)].

The subject application points out (see page 7, lines 12-20; pages 17-19) a number of publicly accessible facilities that include synchrotrons with undulator beamlines, including the Argonne National Laboratory's Advances Photon Source (APS). Other similar facilities are listed in Hendrickson et al., "Phase Determination by the Method of Multiwavelength Anomalous Diffraction (MAD)," Methods in Enzymology, 276:494-523 (1997), which is incorporated in the subject application by reference (page 25, lines 27-29). The subject application also points out (for example, at pages16-20) desirable attributes of the X-ray crystallography apparatus. Thus, one of ordinary skill in the art can practice the claimed invention by following the guidance provided by the subject application.

The practice of some patentable inventions will require great expense. There is no bar against the patenting of inventions that are costly to practice. Moreover, Applicants maintain that the required enabling disclosure of the claimed invention need not correspond to the most economically efficient embodiments.

Further, Applicants maintain that the patent laws and rules do not require each component of the claimed invention to be commercially available for purchase, and the application need not disclose how to mass produce the claimed invention [see Christianson v. Colt Industries Operating Corp., 822 F.2d 1544, 1562 (Fed.Cir. 1987)]. As acknowledged in the October 23, 2001 Office Action, a member of the public can apply for use of the beamlines at, for example, the APS facilities. Thus, the subject application provides a person of ordinary skill in the art with a sufficiently enabling disclosure with regard to performing high-throughput crystallography.

In addition, although the clustering task may be non-trivial, the

Heger paper shows that the task can be accomplished with the bioinformatics tools available in the art. Moreover, the subject application (at page 14) identifies, for example, one well-known bioinformatics tool (i.e. BLAST) which may be used. Other bioinformatics tools are available and also may be used. Thus, Applicants maintain that the subject application provides an enabling disclosure with regard to clustering of proteins into families.

The subject application (at page 23, lines 2-10) also points out that even if crystallization techniques had not yet been developed for a few protein families at the time of filing of the subject application, the claimed invention of the present application can be practiced to determine the structure of target proteins in such families once crystallization techniques for such protein families are developed. The relevant case law establishes that even if some embodiments of a broadly claimed invention are inoperative at the time of filing, the claimed invention may nevertheless be sufficiently enabled for the scope claimed if at least some of the disclosed embodiments are operative [see, for example, Exxon Research and Engineering Company v. United States, 265 F.3d 1371, 1382 (Fed.Cir. 2001); and Atlas Powder Company v. E.I. DuPont De Nemours & Company, 750 F.2d 1569, 1576 (Fed.Cir. 1984)].

Accordingly, Applicants respectfully request that the Examiner reconsider and withdraw the rejection of claims 1-12 under 35 U.S.C. §112, first paragraph.

Rejection under 35 U.S.C. §112, second paragraph

On page 5 of the January 27, 2003 Office Action, claims 1-12 were rejected under 35 U.S.C. §112, second paragraph, as allegedly indefinite for failing to particularly point out and distinctly

claim the subject matter which the applicant regards as the invention.

The Examiner stated that "the claims recite 'information for selected proteins.'" The Examiner also stated that "it is unclear how a protein is selected (what parameters of criteria are used) and how many proteins are selected." The Examiner further stated that "the metes and bounds of this selection cannot be determined."

The Examiner stated that "the claims recite 'homologous sequences.'" The Examiner further stated that "it is unclear what level of homology is required to meet the limitation of the claim."

The Examiner stated that "the claims recite 'a plurality of target proteins which are members of the family.'" The Examiner also stated that "the criteria that define a family are not provided." The Examiner further stated that "it is unclear how a target is selected (what parameters or criteria are used) and how many targets are selected."

The Examiner stated that "the claims recite 'screening products of the synthesis to choose selected synthesized products for processing.'" The Examiner further stated that "the criteria or parameters for the selection are not provided."

The Examiner stated that "the claims recite 'crystallization screens.'" The Examiner further stated that "the metes and bounds of what is screened and what information is used or produced is not set forth or defined."

The Examiner stated that "the claims recite 'suitable specimen

crystals.'" The Examiner also stated that "it is unclear what is required to meet the limitation of 'suitable.'" The Examiner further stated that "the metes and bounds of this term are not provided."

The Examiner stated that "the claims recite updating the database without providing a clear direction as to what is updated." The Examiner questioned "are new entries created and added to the database, fields added, fields edited or annotated, or is something else intended?" The Examiner also questioned "is the updating with respect to each and every target or with respect to the predicted class of compounds (each and every predicted compound?) and/or homology model of predicted protein structures and/or something else?"

The Examiner stated that "Claims 1 and 7 recite 'binding potency using the active sites information corresponding to the target protein.'" The Examiner also stated that "this is confusing as the target may or may not have binding potency (whatever the metes and bounds of this phrase may be which are not defined) nor any active sites (again the metes and bounds of what would meet this limitation is not provided)." The Examiner further stated that "it is unclear exactly what is being modeled in the portions following this phrase." The Examiner stated that "it appears it could be either proteins similar to the target (in the same family) or proteins that bind to or interact with the target."

The Examiner stated that "with respect to claims 1-6, the claimed system does not set forth the relationship of the database, bioinformatics tool, protein synthesis means, protein processing means, crystallization means, X-ray crystallography means, and so forth." The Examiner also stated that "the claim language does not reflect an integrated or turn-key system where the components

are related or linked to each other in some fashion." The Examiner further stated that "as written, the claim appears to be directed to a collection of laboratory equipment or machines." The Examiner stated that "this does not appear to define a system."

The Examiner stated that "with respect to Claims 7-12, the claimed method fails to particularly point out what steps are to be performed and how they are to be performed."

In response, without conceding the correctness of the Examiner's position but solely to advance the prosecution of the present application, Applicants have hereinabove amended claims 1-12. Applicants maintain that the claim amendments do not narrow the scope of the claimed invention, but rather place the claims in better form for examination.

Applicants traverse the rejection with regard to the following.

With regard to the phrase "homologous sequences", Applicants respectfully refer the Examiner to the specification at, for example, page 14, lines 5-9 and 18-23, and page 21, lines 1-6. As pointed out in the subject application, one or more sequence analysis programs known in the art (for example, BLAST) may be used. Further, Applicants submit that one of ordinary skill in the art would also understand that, generally, when a higher level of homology is required, the likelihood of "homologous sequences" being associated with dissimilar corresponding three-dimensional structures is lower, but the size of the families tend to be smaller. Conversely, when a lower level of homology is required, the size of the families tend to be larger, but the likelihood of homologous sequences being associated with dissimilar corresponding three-dimensional structures is higher.

Thus, Applicants maintain that claims 1 and 7 would be clear and unambiguous to one of ordinary skill in the art guided by her/his knowledge of the term "homologous sequence" as used in the art, and further in view of the claims and the specification.

With regard to the phrase "a plurality of target proteins which are members of the family", Applicants maintain that the claim clearly and unambiguously recites that "for each family, members of the family have homologous sequences".

With regard to system claims 1-6, Applicants maintain that the patent laws and rules do not require the claimed invention recited in a system claim to be limited to a turn-key system. Applicants maintain that the elements of the claims have the required functional and/or structural interconnectivity.

With regard to method claims 7-12, Applicants maintain that the steps of systematically organizing [step (a)], clustering [step (b)], synthesizing [step (c)], preparing, purifying and characterizing [step (d)], crystallizing [step (e)], testing [step (f)], performing high-throughput crystallography [step (g)], analyzing [step (h)], developing a homology model [step (i)], and updating the database [step (j)], particularly and distinctly point out the subject matter which Applicants regard as the claimed invention.

Accordingly, Applicants respectfully request that the Examiner reconsider and withdraw the rejection of claims 1-12 under 35 U.S.C. § 112, second paragraph.

In view of the amendments to the claims and remarks hereinabove, Applicants maintain that claims 1-12 are now in condition for allowance. Accordingly, Applicants earnestly solicit the

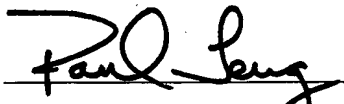
allowance of claims 1-12.

If a telephone interview would be of assistance in advancing prosecution of the present application, Applicant's undersigned attorney invites the Examiner to telephone him at the telephone number provided below.

If a petition for an extension of time is required to make this response timely, this paper should be considered to be such a petition, and the Commissioner is authorized to charge the requisite fees to our Deposit Account No. 03-3125.

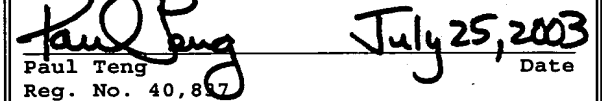
No fee, other than the $465.00 fee for the three-month extension of time, is deemed necessary in connection with the filing of this Amendment. However, if any additional fee is required, authorization is hereby given to charge the amount of any such fee to Deposit Account No. 03-3125.

Respectfully submitted,



John P. White, Reg.No. 28,678
Paul Teng, Reg. No. 40,837
Attorneys for Applicants
Cooper & Dunham, LLP
1185 Avenue of the Americas
New York, New York 10036
(212) 278-0400

I hereby certify that this correspondence is being deposited this date with the U.S. Postal Service with sufficient postage as first class mail in an envelope addressed to: Mail Stop AF, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.



Paul Teng **July 25, 2003**
Reg. No. 40,837 **Date**

1. (Five Times Amended) A system for determining experimentally a plurality of three-dimensional atomic structures, each of which is associated with a corresponding protein, comprising:

a database of sequence information for a first plurality of proteins, and structural information and functional information for [selected] a second plurality of proteins;

at least one bioinformatics tool adapted to use the sequence information, structural information and functional information stored in the database to cluster the first plurality of proteins into a plurality of families, in which, for each family, members of the family have homologous sequences;

protein synthesis means for synthesizing for each family determined by the at least one bioinformatics tool a plurality of target proteins which are members of the family, using information stored in the database corresponding to the target proteins, the protein synthesis means having screening means for screening products of the synthesis to choose selected synthesized products, which are effective as the target proteins, for processing;

protein processing means for preparing, purifying and characterizing each of the selected synthesized products screened through the screening means;

crystallization means for crystallizing the processed synthesized product [against a plurality of crystallization screens] processed by the protein processing means, to produce a plurality of specimen crystals of the target protein, and testing the plurality of specimen crystals for predetermined diffraction characteristics to determine [suitable] the specimen crystals which are suitable for diffraction measurement;

X-ray crystallography means for performing high-throughput crystallography on the specimen crystals [of each target protein]

determined by the crystallization means to be suitable for diffraction measurement, the X-ray crystallography means having diffraction measuring means for measuring for diffraction data the suitable specimen crystals of the target protein, analyzing means for analyzing the diffraction data, means for building an atomic model of the target protein according to an analysis of the diffraction data by the analyzing means, and means for refining the model of the target protein against the diffraction data and storing the refined model in the database;

structure extraction means having means for analyzing the refined model of the target protein using sequence information corresponding to other family members which is stored in the database and structural information corresponding to other proteins which is stored in the database, and means for analyzing the refined model for functional motifs and for surface characteristics [to define active sites and macromolecular contact sites, and means for defining at least one class of compounds predicted to have binding potency using the active sites information corresponding to the target protein]; and

a homology model building tool adapted to use the refined model of the target protein retrieved from the database to develop a homology model of one or more predicted protein structures,

wherein the database is updated using the at least one bioinformatics tool and the developed homology model to link the refined model of the target protein to other databases which store information concerning biological pathways and functional annotation.

7. (Five Times Amended) A process for determining experimentally a plurality of three-dimensional atomic structures, each of which is associated with a corresponding

protein, comprising the steps of:

(a) systematically organizing sequence information for a first plurality of proteins, and structural information and functional information for [selected] a second plurality of proteins into a database;

(b) clustering the plurality of proteins into a plurality of families, in which, for each family, members of the family have homologous sequences, using at least one bioinformatics tool and the sequence information, structural information and functional information stored in the database;

(c) synthesizing for each family determined in step (b) a plurality of target proteins which are members of the family, using information stored in the database corresponding to the plurality of target proteins, and screening products of the synthesis to choose selected synthesized products, which are effective as the target proteins, for processing;

(d) preparing, purifying and characterizing each synthesized product that is chosen in step (c);

(e) crystallizing the processed synthesized product prepared, purified and characterized in step (d) [against a plurality of crystallization screens] to produce a plurality of specimen crystals of the target protein;

(f) testing the plurality of specimen crystals grown in step (e) for predetermined diffraction characteristics to determine [suitable] the specimen crystals [of the target protein] which are suitable for diffraction measurement;

(g) performing high-throughput crystallography, including measuring for diffraction data the specimen crystals determined in step (f) to be suitable for diffraction measurement, building an atomic model of the target protein according to an analysis of the diffraction data, refining the model of the target protein against the diffraction data, and storing the refined model in

the database;

(h) analyzing the refined model, stored in the database in step (g), of the target protein using sequence information corresponding to other family members which is stored in the database and structural information corresponding to other proteins which is stored in the database, and analyzing the refined model of the target protein for functional motifs and for surface characteristics [to define active sites and macromolecular contact sites, and means for defining at least one class of compounds predicted to have binding potency using the active sites information corresponding to the target protein];

(i) developing a homology model of one or more predicted protein structures using computational tools for homology model building and the refined model of the target protein retrieved from the database[,]; and

(j) updating the database by using the at least one bioinformatics tool and the developed homology model [; and] to link the refined model of the target protein to other databases which store information concerning biological pathways and functional annotation,

[(j) performing] wherein steps (f) through [(i)] (j) are repeated for each of the other target proteins.